

## SMU Data Science Review

Volume 1 | Number 3

Article 6

2018

# Minimizing the Perceived Financial Burden Due to Cancer

Hassan Azhar

*Southern Methodist University*, [hazhar@smu.edu](mailto:hazhar@smu.edu)

Zoheb Allam

*Southern Methodist University*, [zallam@smu.edu](mailto:zallam@smu.edu)

Gino Varghese

*Southern Methodist University*, [gjvarghese@smu.edu](mailto:gjvarghese@smu.edu)

Daniel W. Engels

*Southern Methodist University*, [dwe@smu.edu](mailto:dwe@smu.edu)

Sajiny John

*University of Kerala*, [johnsajiny@protonmail.com](mailto:johnsajiny@protonmail.com)

Follow this and additional works at: <https://scholar.smu.edu/datasciencereview>



Part of the [Categorical Data Analysis Commons](#), [Clinical Trials Commons](#), [Statistical Methodology Commons](#), [Statistical Models Commons](#), and the [Statistical Theory Commons](#)

### Recommended Citation

Azhar, Hassan; Allam, Zoheb; Varghese, Gino; Engels, Daniel W.; and John, Sajiny (2018) "Minimizing the Perceived Financial Burden Due to Cancer," *SMU Data Science Review*: Vol. 1 : No. 3 , Article 6.

Available at: <https://scholar.smu.edu/datasciencereview/vol1/iss3/6>

This Article is brought to you for free and open access by SMU Scholar. It has been accepted for inclusion in SMU Data Science Review by an authorized administrator of SMU Scholar. For more information, please visit <http://digitalrepository.smu.edu>.

# Minimizing the Perceived Financial Burden Due to Cancer

Zoheb Allam<sup>1</sup>, Hassan Azhar<sup>1</sup>, Gino Varghese<sup>1</sup>, Daniel W. Engels<sup>1</sup>,  
Sajiny John MBBS.<sup>2</sup>

<sup>1</sup> Master of Science in Data Science, Southern Methodist University,  
Dallas, Texas, USA

{zallam, hazhar, gjvarghese, dwe}@smu.edu

<sup>2</sup> SUT Academy of Medical Sciences, University of Kerala, Kerala, India  
johnsajiny@protonmail.com

**Abstract.** In this paper, we present a regression model that predicts perceived financial burden that a cancer patient experiences in the treatment and management of the disease. Cancer patients do not fully understand the burden associated with the cost of cancer, and their lack of understanding can increase the difficulties associated with living with the disease, in particular coping with the cost. The relationship between demographic characteristics and financial burden were examined in order to better understand the characteristics of a cancer patient and their burden, while all subsets regression and machine learning algorithms were used to determine the best predictors of financial burden. Age, BMI, household income, and breast cancer are the strongest predictors of financial burden, out of which breast cancer patients are most at risk to suffer from high financial burden. Our findings indicate that certain demographic characteristics have a larger impact on perceived financial burden than others, making it possible to target certain demographic groups with increased education in order to aid in the management of the disease.

## 1 Introduction

In the research world one of the topics that is most heavily studied is cancer. The volume of research being conducted is closely related to the rise in incidence. In 2018, there will be an estimated 1,735,350 new cancer cases diagnosed while the total annual cost of cancer care in the US is estimated to reach \$175 billion by 2020, an increase of 40% from 2010 [American Society of Clinical Oncology]. According to the Medical Expenditure Panel Survey from the Agency for Healthcare Research Quality, cancer care costs an average of \$85,201 per patient in 2010-2011, an increase of 40% from 2010 [American Society of Clinical Oncology].

Even with health insurance, cancer patients often find themselves vulnerable to medical debt and bankruptcy. A major contributor is cancer patient's lack of understanding of the costs associated with cancer. Some costs are more obvious such as the cost of chemotherapy, but there are hidden costs as well such as transportation, job loss and travel to and from appointments which

are often not considered. When these hidden costs are ignored, the burden of living with cancer on a day to day basis can increase. Therefore, cancer patients need to be well educated on the costs prior to beginning treatment. In order to better educate cancer patients, first we need to understand the characteristics which could make a cancer patient more susceptible to experiencing financial burden. However, these characteristics are not well understood, and without this understanding reaching the cancer patients in need becomes a difficult task.

Prior research has shown that certain demographic populations have higher cost associated with cancer. However, this approach to understanding financial burden only takes into account the more obvious costs such as direct medical costs and out of pocket expenses. It fails to capture the hidden costs which can be taxing on patients. An efficient method of assessing hidden costs is to measure perceived financial burden. Each cancer patient has different hidden costs, making it difficult to accurately assess the costs. Examining perceived burden allows for a retrospective look into how cancer patient's financial situation was impacted by the diagnosis, and since perception is best measured with self-report questions, you get clear insight into how patients see the burden places upon themselves.

Our results indicate that female participants had higher perceived financial burden than male participants, while Caucasian participants had the least financial burden compared to other ethnicities. Participants who received a recent checkup were also more likely to have higher perceived financial burden, while participants with higher household income were less likely to have higher financial burden. The type of cancer was also important in assessing how participants view their financial burden. Participants with breast cancer had significantly higher perceived financial burden than participants with other forms of cancer.

These results start to illuminate the characteristics of cancer patients who may be susceptible to high financial burden. They also show that not all cancer patients experience financial burden equally, as some types of cancer may place greater burden on patients than others. Therefore, the way we educate cancer patients on the burden they experience cannot be the same for all types of patients and for all types of cancer. A special targeted educational program may be the solution which caters to the specific demographic groups that need the assistance.

To better understand the problem, the paper first goes over what is cancer, the types of cancer, and ways to treat the diagnosis. It then highlights the cost of cancer and the impact and perceived burden. Information on the data file used, the analysis methods, and results are subsequently presented. Finally, the ethics associated with this type of research and the conclusion make up the final sections of the paper.

## 2 Cancer

### 2.1 Introduction

Cancer is a collection of hundreds of related diseases. Some of the body's cells begin to divide and reproduce unchecked and spread into surrounding tissues. Cancer can be made up of trillions of these 'rogue' cells

### 2.2 Types of Cancer

There are many different types of cancer characterized by abnormal cell growth. They are generally identified by the body area they target.

A cancer that forms in the cells of the breasts is known as breast cancer. According to the American Cancer Society, an estimated 22 percent of people live at least five years after being diagnosed with stage 4 breast cancer. Many factors can affect the chances of long-term survival of cancer patients.

A man's prostate produces the seminal fluid that nourishes and transports sperm. Symptoms of cancer in the prostate include difficulty with urination, but in some cases no symptoms present in patients. Approximately 90% of all prostate cancers are detected in the local and regional stages, so the cure rate is very high: Nearly 100% of men diagnosed and treated at this stage will be disease-free after five years. However, prostate cancer is still a deadly disease if not diagnosed early enough. A type of skin cancer begins in the basal cells. Basal cells produce new skin cells as old ones die. Limiting sun exposure and protecting the skin using sunscreens can help prevent these cells from becoming cancerous. This cancer typically appears as a white waxy lump or a brown scaly patch on sun-exposed areas, such as the face and neck. Melanoma occurs when the pigment-producing cells that give color to the skin become cancerous. Symptoms might include a new, unusual growth or a change in an existing mole. Melanomas can occur anywhere on the body. Treatment may involve surgery, radiation, medication, or in some cases chemotherapy. A cancer can also form in the colon or rectum, which is located at the digestive tract's lower end. Early cases can begin as noncancerous polyps. These often have no symptoms but can be detected by screening. Colorectal cancer symptoms depend on the size and location of the cancer. Some common symptoms include changes in bowel habits, changes in stool consistency, blood in the stool, and abdominal pain. Treatment for colorectal cancer also depends on the size, location, and the extent of malignancy. Common treatments include surgical removal, chemotherapy, and radiation therapy.

Lung cancer most often occurs in people who smoke. Two major types of lung cancer are non-small cell lung cancer and small cell lung cancer. Causes of lung cancer include smoking, secondhand smoke, exposure to certain toxins, and family history. Symptoms include cough (often with blood), chest pain, wheezing, and weight loss. These symptoms often don't appear until the cancer is advanced. Leukemia is cancer of blood-forming tissues, including bone marrow. Many types exist such as acute lymphoblastic leukemia, acute myeloid leukemia,

and chronic lymphocytic leukemia. Many patients with slow-growing types of leukemias don't show any symptoms. Rapidly growing types of leukemia may cause symptoms that include fatigue, weight loss, frequent infections, and easy bleeding or bruising. The lymphatic system is the body's disease-fighting network. It includes the lymph nodes, spleen, thymus gland, and bone marrow. The main types of cancer in the lymphatic system are Hodgkin's lymphoma and non-Hodgkin's lymphoma. Symptoms of lymphoma include enlarged lymph nodes, fatigue, and weight loss.

### 2.3 Types of Cancer Treatment

There are many types of cancer treatment. Some people will have only one treatment whereas some will be combinations of treatments such as surgery and radiations.

Surgery is a procedure in which a surgeon removes cancer from the body. During surgery, the tumor and surrounding tissue are removed. Surgery is the oldest type of cancer therapy and remains an effective treatment for many types of cancer today as long as the cancer has not spread too much to other parts of the body. If the cancer has spread, surgery might not be the best treatment and it may be better to have a treatment that reaches all parts of your body, such as chemotherapy, biological therapy or hormone therapy.

Radiation therapy may be used alone or in combination with surgery, chemotherapy, or both. Radiation therapy may also be given with palliative intent. Palliative treatments are not intended to cure. Instead, they relieve symptoms and reduce the pain and suffering caused by cancer. More than half of cancer patients receive some type of radiation therapy. For some cancers, radiation therapy alone is an effective treatment. Other types of cancer respond best to combination of the treatment methods. This may include radiation therapy plus surgery, chemotherapy, or immunotherapy. Chemotherapy (chemo) usually refers to the use of medicines or drugs to treat cancer. The thought of chemotherapy frightens most patients. But knowing what chemotherapy is, how it works, and what to expect can often help calm the fears. It can also give a better sense of control over cancer treatment. Immunotherapy is a type of treatment that utilizes the human body's own immune system to fight cancer. This can be done in two ways. The first involves stimulating the immune system to work harder or smarter to attack cancer cells while the second method involves giving the immune system extra components, such as man-made immune system proteins. A stem cell transplant is also a suitable treatment for some types of cancer. For example, it is effective against leukemia, multiple myeloma and some types of lymphomas. A part of bones called "bone marrow" makes blood cells. Marrow is the soft, spongy tissue inside bones. It contains cells called "hematopoietic" stem cells. These cells can turn into several other types of cells. They can grow into more bone marrow cells. Or they can grow into any other type of blood cell. Certain cancers and other diseases keep hematopoietic stem cells from developing normally. If they are not normal, neither are the blood

cells that they make. A stem cell transplant gives you new stem cells which can make new, healthy blood cells.

## 2.4 Cycles of Cancer Treatment

A series of cycles of treatment is called a course. A treatment course often takes between 3 to 6 months but it can be longer than that. During that time, a patient can probably have between 4 to 8 types of treatment. For example, a person can have both surgery and chemotherapy.

Depending on the combination of treatments, an individual treatment may last from a few months to a few years. Treatment cycles might be weekly or might take 2, 3 or 4 weeks, depending on the drugs and the treatment plan. Some treatment cycles are quite complicated which is defined by the phase of the cancer cells.

## 3 Cost of Cancer

A cancer diagnosis can be expensive and a major source of stress and anxiety. Sometimes, the enormous costs keep people away from completing cancer treatment which increases health risks. Some of the factors that contribute to cost of cancer care are explained below.

The first cost is related to the doctor appointments. This includes payments for care you receive at each doctor visit. Typically, your insurance provider requires you to pay a certain amount from your own pocket, also referred to as a co-payment or co-pay. The co-pay amount is specified by the insurance provider. This fee needs to be paid for each appointment.

You may also have to pay for tests done during your appointment. For example, laboratory tests, such as blood or urine tests. These lab charges generally depend on the difficulty or sophistication of the tests involved.

Another major factor is the cost related to the actual treatment. This includes payments for care you receive during your cancer treatment. For example, each radiation therapy session or chemotherapy infusion.

Medications are also an important part of cancer treatment. Cost of medications include payments for drugs prescribed during your treatment period. For example, chemotherapy or drugs to help manage side effects.

Mobilization costs are sometimes ignored by patients but they should always be considered. They depend on where you are receiving treatment and how you get there. These include but are not limited to the gas, toll, parking fees and taxi, bus, train or airplane fares.

Some people receive treatment far from home, out of choice or necessity. In this case, you may need to pay for lodging. These include costs of household upkeep and family care during cancer treatment. For example, a cancer patient may require childcare or elder care. They also need help with household chores such as cleaning and cooking. Cancer can mentally impact the patients a lot so they also need counselling to help them through this difficult phase of life.

Depending on the severity of the cancer, the patients also need caregiving such as help with cooking, nursing and even simply moving around to get groceries or get to doctor appointments on time.

There are also employment-related, legal, and financial costs. These include costs of professional help with employment, legal, or financial issues. Professionals may help with coping with the loss of wages and consistent income, learning about employment rights under the law, preparing and filing for taxes to account for medical expenses and writing a will.

## 4 Cost Impact

Cancer is one of the leading causes of death and disease in the U.S. The American Cancer Society (ACS) estimates that roughly 1.7 million new cases of cancer will be diagnosed in the U.S. in 2017 and more than 15 million Americans living today have a cancer history.

In 2014 cancer patients paid nearly 4 billion out-of-pocket for cancer treatments. Cancer also represents a significant proportion of total U.S. health care spending.

With more than 200 different types of cancer, there is no one size fits all cancer treatment, but there are several consistent factors that contribute to patients' overall costs for their care.

### 4.1 Cost impact and perceived burden

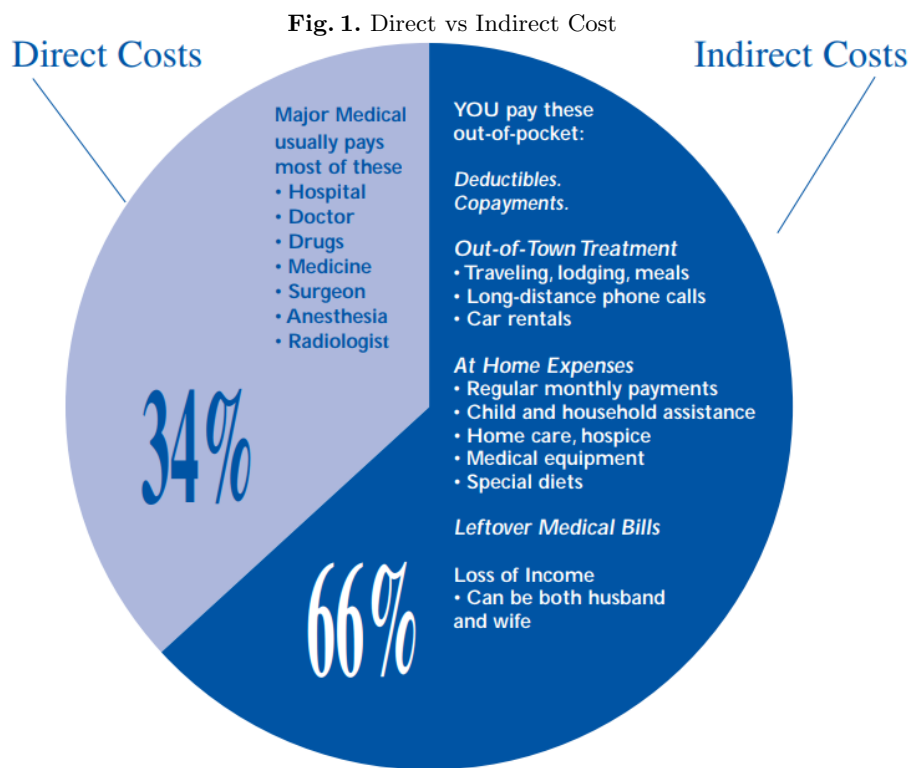
Cancer treatment can be very expensive and lengthy which can take a toll on the impacted individuals. For example, Individuals even with insurance, can be hard hit financially and psychologically by cancer treatment. The costs of treatments including surgery and expensive drugs, as well as non-medical supportive needs, such as child care, can be crippling. These type of financial burden is linked to psychological problems such as depression, mood swings, more worry about cancer recurrence, and lower quality of life. Furthermore, this may lead to foregoing or delaying care or medications which is in turn connected to health-related outcomes and importantly mortality.

### 4.2 Family dynamics

In families with a person diagnosed with cancer, financial issues emerge as a significant concern at a time when these families were already consumed with other challenges. This economic burden can have long-term effects on the financial security, quality of life, and future well-being of the entire family.

### 4.3 Long term side effects

The impact of having cancer does not always end when treatment finishes. The long-term consequences of cancer and its treatment include both physical and mental effects, such as chronic fatigue, sexual difficulties, mental health





problems, pain, urinary and gastrointestinal problems, and lymphoedema (persistent tissue swelling caused by fluid retention). Many of these problems can persist for at least 10 years after treatment and can be significantly worse than those experienced by people without cancer. Certain treatments for cancer also increase the risk of other serious long-term conditions such as heart disease, osteoporosis or a second cancer.

#### **4.4 Short term side effects**

In short term some people living with and beyond cancer experience chronic fatigue which may lead to inability to perform their jobs because they cannot work consistently or efficiently. Some consequences of cancer and its treatment can be reduced with simple interventions, while more complex issues require specialist services.

#### **4.5 Physical impacts**

Some of the examples of the physical impacts include the following: Functional health impairment, such as the inability to undertake labor-intensive tasks, limitations in mobility due to the loss of a limb. Physical health impacts includes coping with pain, nausea, fatigue, reduced hearing and/or vision. Mental health impacts such as lack of sleep, stress, anxiety, depression, personality changes from brain surgery. Cognitive health impairments such as the impact on memory, IQ, and the inability to remain focused.

### **5 Data**

The Health Information National Trends Survey (HINTS) data collection program was developed by the Health Communication and Informatics Research Branch (HCIRB) of the Division of Cancer Control and Population Sciences (DCCPS) as an after effect of the National Cancer Institute's Extraordinary Opportunity in Cancer Communications. This program was designed to help understand and learn more about communication and informatics research that takes place at the NCI. The HINTS data collection program is a national survey which has been administered every other year starting from 2003 and it was created to monitor changes in the rapidly evolving field of health communication. The data was collected in accordance to the privacy act Section 411.285a, 42 USC, and paperwork reduction act of 1995. The information was gathered as an interview process in which participants were asked to complete a set of questions within a duration of 20 to 30 minutes. This data provided insights on changing patterns, needs, information opportunities in health, communications trends, how cancer risks are perceived and what the best practices are. It also assesses cancer information access, usage and provides researchers to test new theories in health communication.

Researchers are using this data to get insights on how adults 18 years and older use different communication methodologies, including but not limited to the Internet, to obtain vital health information for themselves and their family members. Program planners are using this same data to overcome communication barriers across populations by developing effective communication strategies and finally social scientists are using the data to narrow their theories of health communication in the age of technology and to offer new and better recommendations for reducing the burden of cancer throughout the population.

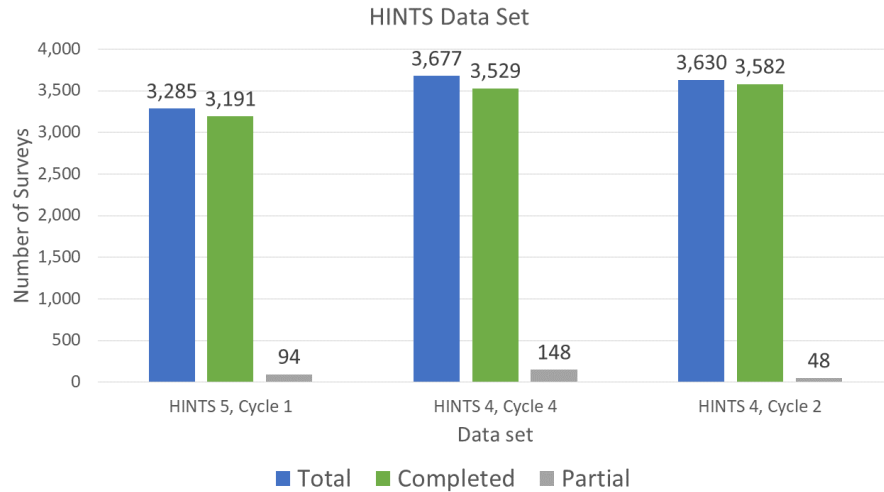
Returned HINTS surveys were reviewed to ensure they were eligible for inclusion in the final dataset, surveys were reviewed for completion and duplication. A survey questionnaire was considered to be complete if at least 80% of Sections A and B were answered, and it's considered to be partially complete if 50% to 79% of the questions were answered in Sections A and B. Handling of duplicate survey were put to the following removal process: If the same respondent returned multiple questionnaires, the first questionnaire was considered complete. If the same respondent returned multiple questionnaires on the same day, the first questionnaire to complete the editing process was considered complete. If a return date was unavailable for questionnaires from the same respondent, the questionnaire with fewer substantive questions omitted was considered complete. If different respondents returned a questionnaire and the ages of household members listed in the roster were in agreement, the questionnaire that complied with the next birthday rule was considered complete.

Every questionnaire that was considered completed, received a full-sample weight. The full-sample weight is used to calculate population and subpopulation estimates, replicate weights were also used to compute standard errors for these estimates, the use of sampling weights is done to ensure valid inferences from the responding sample to the population, correcting for nonresponse and noncoverage biases. Also the replicate weights were calculated using the "delete one" jackknife replication method. For the purpose of this study, we will be ignoring the partially completed survey questionnaire and focusing on the completed ones as shown in the graph.

To avoid bias for attrition from noncontacts, screener nonresponse, weights were assigned to account for all of the stages of selection. These weights were designed to provide approximately unbiased estimators of population. Replicate weights were also provided to allow for consistent variance estimation. The jackknife replication method was used, with  $R = 50$  replicate weights for each survey year for all of the biennial HINTS surveys. A carefully selected portion of the original sample, roughly  $1/50$  of the original sample was deleted and the remaining sample was reweighed as if the complement set was the full sample.

For the purposes of our study, 3 of the HINTS datasets were merged to create one master data file (see Table 1 for more information on the individual data files).

After the datasets were merged, participants who indicated they had never received a diagnosis of cancer were removed from the data file. The final data file consisted of information from 1,595 patients who had received a diagnosis of

**Fig. 2.** HINTS Data Set**Table 1.** Health Information National Trends Survey Data set

HINTS	Cycles	Start	End	N
5	1	January 25, 2017	May 5, 2017	3,191
4	4	August 19, 2014	November 17, 2014	3,529
4	2	October 2012	January 2013	3,582

cancer and 41 different variables. Out of which only 1,359 patients responded to "CancerHurtFinances" variable with valid answers ranging as such (1="Not at all", 2="A little", 3="Some" and 4="A lot"). Some of the variables that were examined included health variables like smoking status, general health indices, and most recent checkup. General demographic variables such as age, gender, ethnicity, etc. were also examined. Finally, different types of cancer diagnosis were examined. Table 2 displays the full list of variables that were used in this study.

## 6 Analysis Methods

All analyses were conducted in SAS and Python. Prior to any model being created, bivariate relationships were tested between perceived financial burden and each of the variables in the data file. Spearman's correlations were conducted to examine the relationship between rank order variables and perceived financial burden. Spearman's coefficient is a statistical measure of the strength of a relationship between two ordinal variables, with coefficient values of .00 - .39 indicating a weak relationship, .40 - .59 indicating a moderate relationship,

and .60 and higher indicating a strong relationship. Independent samples t-tests were also conducted to examine the relationship between dichotomous variables and perceived financial burden, while one-way analysis of variance (ANOVA) was conducted to examine the relationship between variables with more than two levels and perceived financial burden. The independent samples t-Test compared the means of two independent groups in order to determine whether the means from the associated population are significantly different, and the one-way ANOVA determines whether the means from more than two independent groups are significantly different.

To create the best model, machine learning algorithms were utilized. Specifically 3 models were tested and compared to each other: Gaussian naïve bayes, decision tree classifier, and random forest classifier. Machine learning algorithms allow for more accurate results and provides researchers a better chance of identifying important predictors. Therefore, we decided to first run these 3 models to determine which of the three provide the most accurate model. Standard machine learning procedures such as cross validation were used. Of the 1,595 cancer patients we had, 70% of the data was set aside for the training data, 15% for the testing data, and 15% for the final validation data. Testing data is set aside as we want to see how the new model we fit on the training data compared to the original model. Once we have identified the best model, we will have the model identify which are the most important predictors of perceived financial burden.

All subsets regression will then be utilized in order to determine which of the other variables not deemed as most important significant predictors of perceived financial burden are. In statistical modeling, regression analysis is a process for estimating the relationships between a set of independent or predictor variables and a single continuous dependent or outcome variable. Regression allows us to understand how the values of the dependent variable changes when any of the independent variables are varied. All subsets regression is a particularly efficient technique where there is no natural ordering to the explanatory variables.

When conducting all subsets regression, the first step is to identify which combination of predictor variables are best at explaining the outcome variable. In order to do this, a model fit indices must be selected. Four common indices are utilized, Adjusted R-Square, R-Square, AIC, and SSE. For the purposes of this study, adjusted R-square was used to determine the best model. The adjusted R-square tells you the amount of variance explained in the outcome variable by the predictor variables. Once the best model was identified, a multiple linear regression was conducted to determine which of the predictor variables identified in the prior step a significant predictor of perceived financial burden.

## 7 Results

Preliminary analyses aimed at identifying relationships between demographic characteristics and perceived financial burden revealed several significant findings. Female participants had significantly higher perceived financial burden than male participants as shown in Figure. 2. This could be attributed to the difference in

wages between males and females, as studies have indicated that women working full time in the US are typically paid just 80% of what men were paid [1]. Our results also indicated that Caucasian participants had the least financial burden compared to other ethnicities, while American Indians/Alaskan natives had the highest financial burden (also shown in Figure X below). This could also be explained by earning gap between Caucasians and people from different ethnicities. White families earn nearly 10 times more than African American families [2], and the gap is even larger when considering other minorities. Finally, preliminary findings also identified differences in participant's financial burden and the type of cancer they had. Participants with breast cancer had significantly higher perceived financial burden than participants with other forms of cancer. In 2010, the cost of treating breast cancer was approximately \$16.5 billion, higher than any other type of cancer [3]. This burden placed on breast cancer patients could significantly impact the perceived burden as well.

The main regression model indicated several additional demographic predictors of perceived financial burden. Participants who received a recent checkup were more likely to have higher perceived financial burden. Conceptually this makes sense as patients are charged after every visit, therefore the financial burden they experience would naturally be higher after a checkup. General health was also significantly related to financial burden, healthier participants were more likely to have higher perceived financial burden. This finding is contrary to the conventional belief that healthier individuals are less likely to visit the doctor's office, thereby bringing down their financial burden. One possible explanation for this finding is that the relationship between psychosocial stressors and disease is affected by the nature, number, and persistence of the stressors [4]. Healthier patients may feel high levels of stress, which in turn impacts their perceived financial burden.

Age was also a significant demographic predictor, younger people were more likely to have higher perceived financial burden. In general older participants are financially better off than younger participants as they have been working for a longer period of time, so this findings is in line with the belief that younger participants have fewer financial resources than older participants. Finally, participants with lower household income were more likely to have higher perceived financial burden. This finding also ties into the fact that participants who have more wealth and financial resources can cope with financial burden better than participants with low income.

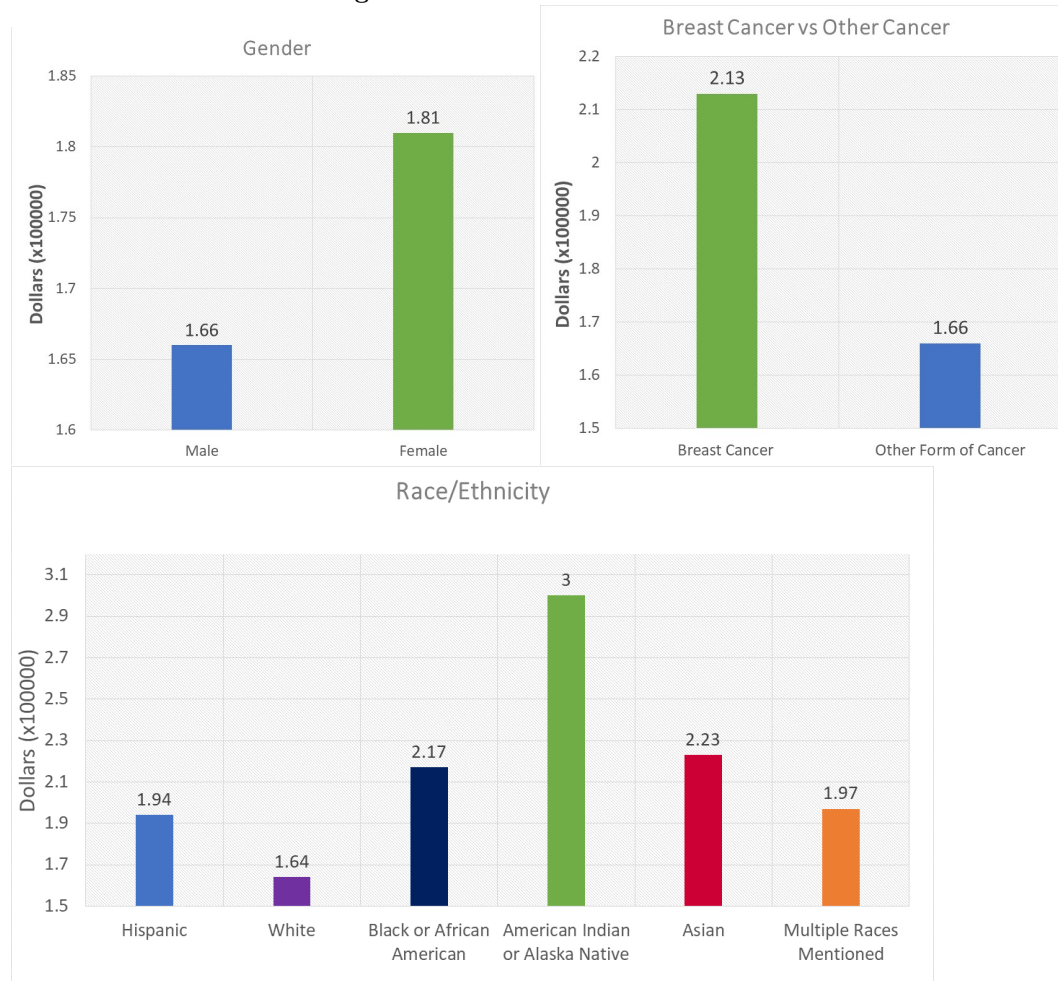
In addition to demographic predictors, the regression model also included different types of cancer to determine if the type of cancer is significantly related to perceived financial burden. Our results revealed that the type of cancer does matter, as not all types of cancer were significantly related to financial burden. Bone, cervical, colon, pancreatic, renal, and stomach cancers were not significant predictors, indicating that patients with these ailments did not show significant differences in their financial burden compared to patients with other types of cancer. Colon and pancreatic cancer are two of the deadliest forms of cancer, and even an early diagnosis does not aid in stopping the spread The 10 Deadliest

Cancers and Why There's No Cure. Therefore, since many patients don't live for a long time with the ailment their cost will be lower, and so our model was unable to detect a significant relationship with financial burden.

There were several types of cancer that were related to perceived financial burden. Bladder, breast, endometrial, head and neck, hodgkins, leukemia, lung, non-hodgkin, pharyngeal, prostate, and rectal cancers were all significant predictors. Breast cancer as mentioned earlier is one of the most expensive forms of cancer, so you would expect to see a relationship with perceived financial burden. Breast cancer is also in the top 5 when it comes to survival rates, with only about 10% of patients dying from the disease within 5 years of receiving the diagnosis Cancer Survival Rate Statistics by Type of Cancer. Prostate, (.8%), bladder (22.1%), and rectal (33.5%) are also in the top 10 when analyzing the survival rate. This relationship with survival rate could help explain why these types of cancers were significantly related to financial burden, as the longer a patient survives with these cancers the higher the cost, which in turn increases the financial burden.

## 8 Ethics

As the cost of research increases, more researchers are turning to surveys for their data collection purposes as they are relatively inexpensive [5]. The dataset we used also came from a survey which was used to describe the characteristics of a larger population but made it affordable for the researchers to access the population. However, when doing survey research there are several ethical components that must be considered. First and foremost, the survey needs to be taken on a voluntary basis by participants [6]. The people who receive this survey need to be told that participation in the survey is completely voluntary. In order to make taking the survey appealing to participants, it is ethical to incentivize participants with monetary values as long as they also understand what the reward will be for completing the survey. Participants also need be informed about the risks, benefits, and facts before they make the decision to participate. Falsely leading participants to take something that could lead to emotional damage takes away participants right to choose whether or not this survey is right for them. Also if you know your survey could be detrimental to participants, as a researcher you should strive to word the questions in such a way that it minimized potential harm to participants. As researchers who use surveys to understand the behavior or opinions of our participants we should attempt to avoid introducing bias into the questions that lead participants to answer a certain way. Not only do you not want participants to be influenced in how they respond to questions, you also do not want the questions themselves or the ordering of the questions to impact how participants respond. Finally, the most important aspect of survey research is to maintain the privacy and confidentiality of survey participants. In most cases surveys won't ask for any information that can identify you, and if they do those questions will be asked on a separate survey so identifying information cannot be linked back to your responses. The data stored from the

**Fig. 3.** Data visualization

surveys should also be kept in a secure location where information cannot be hacked into or leaked. Participants who take the survey are taking it with the faith that the data or any personal information won't be used in an inappropriate fashion and allowing someone to access the data who shouldn't may allow for identity theft etc. The HINTS survey met the requirements for all of the above points, and since the researchers have conducted this survey multiple times the process put into place ensures that the data is secure, participants take the survey of their own free will, consent is received prior to any questions being shown, and the questions are worded to minimize harm to participants.

## 9 Conclusion

While tackling cancer, most cancer patients, might overlook cost related burdens associated with cancer, and their lack of oversight on this burden can increase the difficulties associated with living with the disease. Our model guides patients regarding perceived financial burden due to cancer and minimize difficulties living with the disease. Our findings also indicate that certain demographic characteristics have a larger impact on perceived financial burden. Age, BMI, Time To Moderate Exercise, Education, household income, and breast cancer are the strongest predictors of perceived financial burden. In short cancer is a journey that financially impacts you for the rest of your life.

## 10 Appendix



**Table 2.** Cookbook

Variable Name	Variable Label
MailHHAdults	2. Including yourself, how many people age 18 or older live in this household?.
MostRecentCheckup	C3. About how long has it been since you last visited a doctor for a routine checkup? A routine checkup is a general physical exam, not an exam for a specific injury, illness, or condition.
HealthInsurance	C2. Do you have any kind of health care coverage, including health insurance, prepaid plans such as HMOs or government plans such as Medicare?
GeneralHealth	F1. (E1.) In general, would you say your health is...
TimesModerateExercise	H1. (F5.) In a typical week, how many days do you do any physical activity of at least moderate intensity?
smokeStat	Smoking Status
EverHadCancer	M1. (J1.) Have you ever been diagnosed as having cancer?
TimeSinceDX	How long ago were you diagnosed with cancer?
Age	O1. (L1.) What is your age?
OccupationStatus	O2. (L2.) What is your current occupational status?
MaritalStatus	O5. (L5.) What is your marital status?
Education	O6. (L6.) What is the highest grade or level of schooling you completed?
BornInUSA	O7. (L7.) Were you born in the United States?
BMI	Body Mass Index (Weight*703)/(Height in inches**2)
RaceEthn	Race/Ethnicity.
RentOrOwn	O15. (L15.) Do you currently rent or own your home?
HHInc	What is your combined annual household income?
GenderC	K1. (H1.) Are you male or female?
CancerHurtFinances	M10. Looking back, since the time you were first diagnosed with cancer, how much, if at all, has cancer and its treatment hurt your financial situation?
CaBladder	M2. (J2.) What type of cancer did you have? Bladder cancer
CaBone	M2. (J2.) What type of cancer did you have? Bone cancer
CaBreast	M2. (J2.) What type of cancer did you have? Breast cancer
CaCervical	M2. (J2.) What type of cancer did you have? Cervical cancer (cancer of the cervix)
CaColon	M2. (J2.) What type of cancer did you have? Colon cancer
CaEndometrial	M2. (J2.) What type of cancer did you have? Endometrial cancer (cancer of the uterus)
CaHeadNeck	M2. (J2.) What type of cancer did you have? Head and neck cancer
CaHodgkins	M2. (J2.) What type of cancer did you have? Hodgkin's lymphoma
CaLeukemia	M2. (J2.) What type of cancer did you have? Leukemia/Blood cancer
CaLiver	M2. (J2.) What type of cancer did you have? Liver cancer
CaLung	M2. (J2.) What type of cancer did you have? Lung cancer
CaMelanoma	M2. (J2.) What type of cancer did you have? Melanoma
CaNonHodgkin	M2. (J2.) What type of cancer did you have? Non-Hodgkin lymphoma
CaOral	M2. (J2.) What type of cancer did you have? Oral cancer
CaOvarian	M2. (J2.) What type of cancer did you have? Ovarian cancer
CaPancreatic	M2. (J2.) What type of cancer did you have? Pancreatic cancer
CaPharyngeal	M2. (J2.) What type of cancer did you have? Pharyngeal (throat) cancer
CaProstate	M2. (J2.) What type of cancer did you have? Prostate cancer
CaRectal	M2. (J2.) What type of cancer did you have? Rectal cancer
CaRenal	M2. (J2.) What type of cancer did you have? Renal (kidney) cancer
CaSkin	M2. (J2.) What type of cancer did you have? Skin cancer, non-melanoma
CaStomach	M2. (J2.) What type of cancer did you have? Stomach cancer

**Table 3.** Results

Variable	Parameter Estimate	Standard Error	t Value	Pr > t
Intercept	13.89	3.66	3.80	0.00
MailHHAdults	0.05	0.05	1.09	0.27
MostRecentCheckup	0.06	0.05	1.29	0.20
GeneralHealth	0.11	0.04	2.62	0.01
TimesModerateExercise	-0.02	0.02	-1.23	0.22
TimeSinceDX	-0.04	0.04	-1.07	0.28
Age	-0.02	0.00	-5.71	<.0001
RaceEthn	0.10	0.04	2.71	0.01
HHInc	-0.12	0.03	-4.05	<.0001
CaBladder	-0.78	0.28	-2.78	0.01
CaBone	1.26	0.91	1.38	0.17
CaBreast	-0.67	0.10	-6.54	<.0001
CaCervical	0.22	0.14	1.60	0.11
CaColon	-0.24	0.16	-1.48	0.14
CaEndometrial	-0.52	0.20	-2.53	0.01
CaHeadNeck	-0.56	0.27	-2.08	0.04
CaHodgkins	-0.65	0.31	-2.07	0.04
CaLeukemia	-0.52	0.27	-1.95	0.05
CaLung	-0.74	0.31	-2.37	0.02
CaNonHodgkin	-0.79	0.24	-3.29	0.00
CaPancreatic	1.12	0.91	1.23	0.22
CaPharyngeal	-0.97	0.46	-2.10	0.04
CaProstate	-0.32	0.12	-2.71	0.01
CaRectal	-0.51	0.29	-1.72	0.09
CaRenal	-0.31	0.23	-1.33	0.18
CaStomach	-0.69	0.46	-1.51	0.13

## References

1. et. al, L.: Financial and family burden associated with cancer treatment in ontario, canada. *Support Care Cancer* **14** (2006) 1077–1085
2. et. al, C.: Self-reported financial burden and satisfaction with care among patientst with cancer. *The Oncologist* **4** (2014) 414–420
3. et. al, S.: [the financial toxicity of cancer treatment: a pilot study assessing out-of-pocket expenses and the insured cancer patient’s experience.]. *Oncologist* **18** (2013) 381–390
4. Arozullah, A., Calhoun, E., Wolf, M., K Finley, D., Fitzner, K., A Heckinger, E., S Gorby, N., Schumock, G., Bennett, C.: The Financial Burden of Cancer: Estimates from a Study of Insured Women with Breast Cancer
5. ASCO: Managing the cost of cancer care (2018) Online; accessed 20 January. 2018.
6. of Colorado Cancer Center, U.: Cancer by the number (2015) Online; accessed 18 January. 2018.
7. et. al, Y.: Economic burden of cancer in the us: Estimates, projections, and future research. *Cancer Epidemiol Biomarkers Prev.* **10** (2006) 1151–1158
8. et. al, K.: Mortality, morbidity, and cost associated with febrile neutropenia in adult cancer patients. *Support Care Cancer* **14** (2006) 2258–2256
9. et. al, S.: Stress and health: Psychological, behavioral, and biological determinant. *Annual Review in Clinical Psychology* **1** (2008) 607–628
10. Australia, C.C.: Estimating the economic cost of cancer
11. Network, A.C.S.C.A.: The costs of cancer addressing patient costs
12. Sharing, J.I.: Paperwork reduction act (pra), 44 u.s.c. § 3501 et seq. (1980) Online; accessed 20 January 2018.
13. Westat: Health information national trends survey (2017) Online; accessed 20 October. 2017.
14. Economics, A.: Estimating the economic costs of cancer (2011) Online; accessed 22 October. 2017.
15. Network, C.A.: The cost of cancer (2017) Online; accessed 21 January. 2018.
16. Society, A.C.: Direct vs. indirect expenses (1998) Online; accessed 17 January. 2018.
17. Support, M.C.: Long-term consequences of cancer and its treatment by macmillan (2013) Online; accessed 15 January. 2018.
18. Miedema B, Easley J, F.P.H.R.M.M.: The economic impact on families when a child is diagnosed with cancer (2008) Online; accessed 18 April. 2018.
19. Staff, M.C.: Lymphoma (2018) Online; accessed 15 January. 2018.
20. of Clinical Oncology, A.S.: What is cancer surgery? (2018) Online; accessed 20 April. 2018.